



# Face image super-resolution using 2D CCA

Le An\*, Bir Bhanu

Center for Research in Intelligent Systems, University of California, Riverside, CA 92521, USA



## ARTICLE INFO

### Article history:

Received 25 June 2013

Received in revised form

30 September 2013

Accepted 3 October 2013

Available online 11 October 2013

### Keywords:

Super-resolution

Subspace

1D CCA

2D CCA

Face

Learning

## ABSTRACT

In this paper a face super-resolution method using two-dimensional canonical correlation analysis (2D CCA) is presented. A detail compensation step is followed to add high-frequency components to the reconstructed high-resolution face. Unlike most of the previous researches on face super-resolution algorithms that first transform the images into vectors, in our approach the relationship between the high-resolution and the low-resolution face image are maintained in their original 2D representation. In addition, rather than approximating the entire face, different parts of a face image are super-resolved separately to better preserve the local structure. The proposed method is compared with various *state-of-the-art* super-resolution algorithms using multiple evaluation criteria including face recognition performance. Results on publicly available datasets show that the proposed method super-resolves high quality face images which are very close to the ground-truth and performance gain is not dataset dependent. The method is very efficient in both the training and testing phases compared to the other approaches.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Face is commonly used to recognize humans. In real-world applications, such as video surveillance, detected faces are often of low-resolution, which makes the recognition task difficult. Face image super-resolution, also referred to as face hallucination, is a natural solution to solve this problem. Although in some work super-resolution and recognition are handled simultaneously without generating high-resolution images [1], it is still desirable to obtain a super-resolved face image from low-resolution feed in the case where examination or validation by human is required.

The low-resolution (LR) images can be considered as being generated by the imaging process where the original high-resolution (HR) images undergo blurring and down-sampling [2]. Usually noise is introduced to further degrade

the image quality. The purpose of super-resolution is to reverse this imaging process in order to recover the high-resolution images from the low-resolution observations.

In the past several decades, various super-resolution methods have been proposed. Based on the input, those methods can be categorized into two classes. The methods in the first class take advantage of multiple images of the same scene and reconstruct HR images by aggregating information from all the LR images via motion estimation or registration [3–5]. However, these methods strongly rely on accurate motion information. The methods for super-resolution in the second class are based on a single image. This class of methods has received a lot of attention recently [6–8]. Some learning based methods try to model the relationship between the LR face images and the HR face images [9,10]. An example of a learning-based approach is [11]. Recently, a joint learning approach is proposed by Gao et al. [12] in which two projection matrices are trained simultaneously and the original LR and HR feature spaces are mapped onto a unified feature subspace. This produces improved results compared to the

\* Corresponding author. Tel.: +1 9518273954.

E-mail addresses: [lan004@ucr.edu](mailto:lan004@ucr.edu) (L. An), [bhanu@cris.ucr.edu](mailto:bhanu@cris.ucr.edu) (B. Bhanu).

neighbor-embedding based methods. A sparse neighbor selection scheme is proposed in [13] for image SR and achieves *state-of-the-art* results.

Beyond the generic super-resolution algorithms, specific approaches for certain kind of images such as face images have been proposed [14–17]. Due to the highly structured shape of a human face, more accurate face super-resolution can be achieved by learning this structural information from an appropriate training dataset. In [18], a semi-coupled dictionary learning model is proposed with application to super-resolution and face synthesis. In [15], a two-step face hallucination approach is developed by first globally modeling the face with a Gaussian assumption and then locally refining the face using path-based nonparametric Markov random field (MRF). Inspired by this work, a number of two-step face hallucination approaches have been proposed [14,16,17].

As indicated in [19], the downsampling process maintains the structure in the image manifolds. Furthermore, previous research has shown that face images reside on a non-linear manifold which is linear and smooth locally and it is commonly assumed that the manifolds of LR face images and HR face images have a similar local structure [20]. As a consequence, in recent years, manifold learning has been explored by researchers to hallucinate face images under the assumption that the manifolds of LR and HR images have similar local neighborhood structures [21,22,17]. These methods directly work in the subspaces of LR and HR images using standard subspace techniques such as the Principal Component Analysis (PCA). Compared to performing reconstruction in the original input feature space, the reconstruction is more meaningful and reliable in the subspaces. In a different manner, Ma et al. [23] bypassed the necessity for subspace learning by reconstructing each small patch of a face image separately. A fast face super-resolution method was proposed recently by substituting the nonlinear mapping with multiple local linear transformations [24]. Instead of super-resolving in the image domain, a feature-domain super-resolution framework is proposed in [25] for face recognition. In the following we describe the motivation and contributions of the proposed approach.

### 1.1. Motivation and contributions

Among the manifold learning approaches, canonical correlation analysis has been widely adopted in recent years. Canonical correlation analysis (denoted as 1D CCA in this paper) was first introduced in [26]. It is a multivariate statistical model to analyze the correlation between two sets of variables. 1D CCA finds linear combinations of the variables in each set that have maximum correlations. The model that 1D CCA delivers is a high-dimensional relationship between two sets of variables with a few pairs of canonical variables. There are different generalizations of 1D CCA such as kernel CCA [27], locality preserving CCA (LPCCA) [28], and neural network based CCA [29].

There have been many applications utilizing 1D CCA and its variants. In [30] a 2D–3D face matching method is proposed using 1D CCA to capture the relationship between 2D face images and 3D face data for recognition. Recently

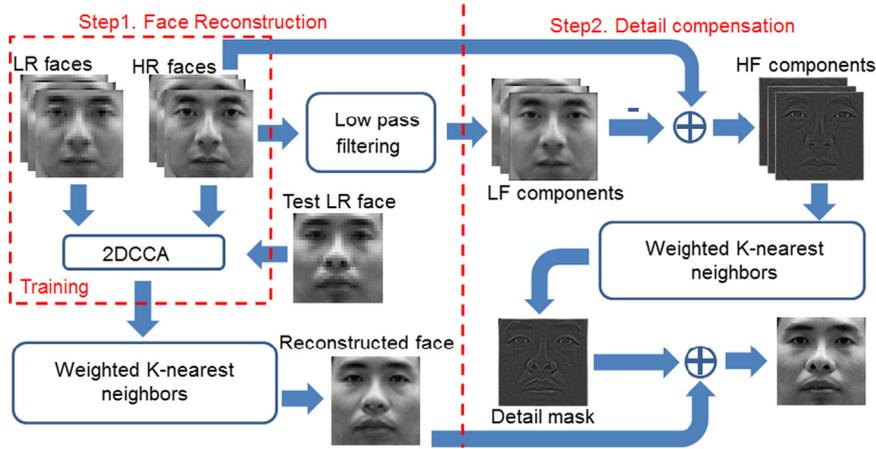
kernelized CCA has been applied in facial expression recognition [31]. Xu et al. [32] proposed a multimodal recognition scheme with ear and profile face using kernel CCA. Beyond the face domain, 1D CCA is also broadly applied. In [33] data fusion and group analysis of biomedical data are performed using 1D CCA. In [34] 1D CCA is used to analyze remotely sensed data in a geographic information system. The handwritten character recognition is also formulated in a framework with 1D CCA [35]. Recently, Li et al. [36] used 1D CCA to maximize the intra-individual correlations for face recognition at different poses.

In terms of super-resolution, Huang et al. [14] proposed a face hallucination method based on 1D CCA to determine a coherent subspace in which the correlation between the LR and HR images is maximized. The face images are first vectorized and projected to PCA subspace, then 1D CCA is applied to enhance the correlation of the HR and LR image projections. This approach tries to find the consistency between the LR and HR face images and is able to generate more realistic face images compared to the previous work in [15,17]. Recently a vehicle logo super-resolution method using 1D CCA is reported in [37] to improve the vehicle recognition accuracy.

However, the 1D CCA was not designed specifically for the image data. To fit the image data into 1D CCA formulation, the image has to be first converted into a 1D vector. On the other hand, an image is inherently represented in a 2D matrix. The appearance of an image becomes obsolete when reshaped into a vector. To tackle this problem, two-dimensional CCA has been proposed and it is specifically suitable for image analysis [38]. 2D CCA is formulated in such a manner that it takes two sets of images and explores their relations directly without the necessity to first vectorize each image.

For face image super-resolution (SR), as a common routine for the data representation in manifold based SR methods, the face images are first reshaped into vectors and then super-resolution is performed. The reshaped vectors have large dimensions. For computational feasibility, PCA is applied [14,17]. However, in this reshaping process the intrinsic 2D spatial structure information of face is erased. In this paper, inspired by the 2D CCA techniques [38], a two-step 2D CCA based face super-resolution approach is developed that can preserve the intrinsic 2D spatial structure of face images in the super-resolution process. In the *first step* the HR face is reconstructed. Since the reconstructed face is not rich in facial details, we apply a high-frequency detail mask to the reconstructed faces in the *second step*. Fig. 1 shows the system diagram of the proposed approach.

More specifically, during the training in the *first step*, the learned projection matrices by 2D CCA are able to project the HR face images and LR face images into a subspace where their correlation is maximized. When a testing LR face image is provided, the optimal combination of its  $K$  nearest neighbors in the LR training set is found in this subspace. Due to the structural similarity between the HR and LR face images, we are then able to reconstruct a HR face image using the same  $K$  nearest neighbors in the HR training set. In the *second step*, a high-frequency detail mask is generated using the neighborhood derived in the



**Fig. 1.** The system diagram of the proposed face super-resolution algorithm. The training set contains HR and LR face images and they are projected to a coherent subspace by 2D CCA in which the correlation between the HR face images and the LR face images is maximized. Given an input LR image, it is first projected into this subspace using the projection matrices obtained in the training step.  $K$  weighted nearest neighbors in the LR training set are found that reconstruct the input LR face with minimum error. The same neighborhood is applied to the HR training set to generate the super-resolved image. A detail compensation step is followed by reconstructing a high-frequency mask from the high-frequency components. The final output is the sum of the reconstructed face image and the detail mask.

first step and added to the reconstructed face image to yield the final output HR image. Inspired by the locality based method [39], we process different parts of a face independently instead of a holistic approach to further improve the output quality.

Compared to the previous work for face image SR including the 1D CCA based method in [14], the contributions of this paper are as follows:

1. This is the first paper that explores 2D CCA for face image super-resolution. To the authors' best knowledge this has not been done before using 2D CCA. The proposed method demonstrates superior performance compared to the *state-of-the-art* super-resolution methods [10,14,23,8] (see Fig. 7).
2. The proposed 2D CCA super-resolution algorithm works directly on the original 2D face image representation. The method is computationally efficient and it achieves the best performance compared to the other methods [10,14,23,8] (see Table 3).
3. Thorough experiments are conducted to validate the approach, both quantitatively and qualitatively, using comprehensive metrics including reference based metrics (PSNR, SSIM, SVD [40]) and non-reference based metric (DM [41]). Cross-dataset validation is also performed (see Fig. 8). Results from the experiments show that the approach is not datasets or image dependent, which is crucial from practical considerations (see Fig. 7). In addition, a recognition task using the super-resolved images by the proposed method leads to the highest recognition accuracy compared to the other methods (see Table 2).

In the rest of this paper, Section 2 provides mathematics for 1D CCA and 2D CCA. Section 3 presents the proposed face super-resolution algorithm. The experimental results and comparisons are given in Section 4. Finally, Section 5 concludes this paper.

## 2. 1D and 2D CCA

### 2.1. 1D CCA formulation

1D CCA was first introduced in [26]. 1D CCA finds basis for two sets of random variables such that the correlation between the projections of these two sets of random variables is maximized. Given two centered (zero mean) datasets,  $X = \{x_i \in \mathbb{R}^m, i = 1, 2, \dots, N\}$  and  $Y = \{y_i \in \mathbb{R}^n, i = 1, 2, \dots, N\}$ , 1D CCA aims at obtaining two basis vectors  $W_X \in \mathbb{R}^m$  and  $W_Y \in \mathbb{R}^n$  such that the correlation coefficient  $\rho$  of  $W_X^T X$  and  $W_Y^T Y$  is maximized. The objective function to be maximized is given by

$$\begin{aligned} \rho &= \frac{\text{Cov}(W_X^T X, W_Y^T Y)}{\sqrt{\text{Var}(W_X^T X)} \sqrt{\text{Var}(W_Y^T Y)}} \\ &= \frac{W_X^T C_{XY} W_Y}{\sqrt{W_X^T C_{XX} W_X} \sqrt{W_Y^T C_{YY} W_Y}} \end{aligned} \quad (1)$$

where  $C_{XX}$  and  $C_{YY}$  are the autocovariance matrices of  $X$  and  $Y$ .  $C_{XY}$  denotes the covariance matrix of  $X$  and  $Y$ .

Equivalently, the 1D CCA can be formulated as a constrained optimization problem by

$$\text{argmax}_{W_X, W_Y} W_X^T C_{XY} W_Y \quad (2)$$

subject to  $W_X^T C_{XX} W_X = 1$  and  $W_Y^T C_{YY} W_Y = 1$ .

### 2.2. 2D CCA formulation

For some data types, such as image, the data representation is inherently two dimensions. Thus, it is desirable to analyze data in the original 2D space without reshaping the data into 1D vectors. Motivated by 2D Principal Component Analysis (2D PCA) [42], 2D CCA was recently developed in [38]. Given two centered datasets,  $X = \{x_i \in \mathbb{R}^{m_x \times n_x}, i = 1, 2, \dots, N\}$  and  $Y = \{y_i \in \mathbb{R}^{m_y \times n_y}, i = 1, 2, \dots, N\}$ , 2D CCA seeks two left projection matrices

$L_X \in \mathbb{R}^{m_x \times d_1}$  and  $L_Y \in \mathbb{R}^{m_y \times d_1}$  and two right projection matrices  $R_X \in \mathbb{R}^{n_x \times d_2}$  and  $R_Y \in \mathbb{R}^{n_y \times d_2}$  such that the correlation coefficient  $\rho$  between the two projected datasets  $L_X^T X R_X$  and  $L_Y^T Y R_Y$  is maximized.  $\rho$  is given by

$$\rho = \frac{\text{Cov}(L_X^T X R_X, L_Y^T Y R_Y)}{\sqrt{\text{Var}(L_X^T X R_X)} \sqrt{\text{Var}(L_Y^T Y R_Y)}} \quad (3)$$

$\rho$  can be written in two parts as

$$\rho_L = \frac{L_X^T C_{XY}^R L_Y}{\sqrt{L_X^T C_{XX}^R L_X} \sqrt{L_Y^T C_{YY}^R L_Y}} \quad (4)$$

$$\rho_R = \frac{R_X^T C_{XY}^L R_Y}{\sqrt{R_X^T C_{XX}^L R_X} \sqrt{R_Y^T C_{YY}^L R_Y}} \quad (5)$$

where  $C_{XX}^R$  is the autocovariance matrix of  $X R_X$ ,  $C_{YY}^R$  is the autocovariance matrix of  $Y R_Y$ , and  $C_{XY}^R$  is the covariance matrix of  $X R_X$  and  $Y R_Y$ . Similarly,  $C_{XX}^L$  is the covariance matrix of  $L_X^T X$ ,  $C_{YY}^L$  is the covariance matrix of  $L_Y^T Y$ , and  $C_{XY}^L$  is the covariance matrix of  $L_X^T X$  and  $L_Y^T Y$ . The equivalent constrained problem for 2D CCA is

$$\arg\max_{L_X, L_Y, R_X, R_Y} \text{Cov}(L_X^T X R_X, L_Y^T Y R_Y) \quad (6)$$

subject to  $\text{Var}(L_X^T X R_X) = 1$  and  $\text{Var}(L_Y^T Y R_Y) = 1$ .

### 2.3. Difference in solving 1D CCA and 2D CCA

Note that the optimization required for 1D CCA in (2) and 2D CCA in (6) is different. Using the Lagrange multiplier, the solution of the optimization problem for 1D CCA is equivalent to the solution of the following generalized eigenvalue problems:

$$\begin{aligned} C_{XY} W_Y &= \lambda C_{XX} W_X \\ C_{YX} W_X &= \lambda C_{YY} W_Y \end{aligned} \quad (7)$$

where  $C_{YX} = C_{XY}^T$ . However, the generalized eigenvalue problem for 2D CCA is different, it involves the following two sets of equations:

$$\begin{aligned} C_{XY}^R L_Y &= \lambda C_{XX}^R L_X \\ C_{YX}^R L_X &= \lambda C_{YY}^R L_Y \end{aligned} \quad (8)$$

$$\begin{aligned} C_{XY}^L R_Y &= \lambda C_{XX}^L R_X \\ C_{YX}^L R_X &= \lambda C_{YY}^L R_Y \end{aligned} \quad (9)$$

The projection matrices  $L_X$ ,  $L_Y$  and  $R_X$ ,  $R_Y$  are solved in an iterative manner. At each iteration, to obtain the updated  $L_X$  and  $L_Y$ ,  $R_X$  and  $R_Y$  are fixed, and  $L_X$  and  $L_Y$  are obtained by computing the  $d_1$  largest generalized eigenvectors in (8). Similarly, to obtain the updated  $R_X$  and  $R_Y$ ,  $L_X$  and  $L_Y$  are fixed, and  $R_X$  and  $R_Y$  are obtained by computing the  $d_2$  largest generalized eigenvectors in (9). This process continues until convergence when the updates from the last iteration to the current iteration become very small. In our experiments,  $L_X$ ,  $L_Y$  and  $R_X$ ,  $R_Y$  converge in a few iterations.

### 3. 2D CCA for face super-resolution

The proposed face super-resolution approach consists of two steps: the *first step* is face reconstruction and the *second step* is detail compensation which further refines a face reconstructed in the first step since a reconstructed face through manifold learning often does not contain sufficient details.

#### 3.1. Face reconstruction

There are two key parts for face reconstruction using 2D CCA: training and reconstruction. During the training, a 2D CCA model is learned. For reconstruction, the learned model is used to construct HR faces from the LR input. It is to be noted that there exists no publication for 2D CCA for face super-resolution and face is an important structure with enormous number of applications. Further, we will see in Section 4 that a 2D CCA based approach provides the best performance compared to almost all the recently published papers on super-resolution.

##### 3.1.1. Training

In the training, 2D CCA is applied to find the left and right projection matrices that project the HR and LR images into a subspace in which the correlation between the projections is maximized. Given the HR training set  $X = \{x_i \in \mathbb{R}^{m_x \times n_x}, i = 1, 2, \dots, N\}$  and the corresponding LR training set  $Y = \{y_i \in \mathbb{R}^{m_y \times n_y}, i = 1, 2, \dots, N\}$ , the mean faces  $\mu_X$  and  $\mu_Y$  are subtracted to obtain the centered datasets  $\hat{X}$  and  $\hat{Y}$ , respectively.

The left transforms  $L_{\hat{X}}$  and  $L_{\hat{Y}}$  and the right transforms  $R_{\hat{X}}$  and  $R_{\hat{Y}}$  are obtained by maximizing

$$\rho = \frac{\text{Cov}(L_{\hat{X}}^T \hat{X} R_{\hat{X}}, L_{\hat{Y}}^T \hat{Y} R_{\hat{Y}})}{\sqrt{\text{Var}(L_{\hat{X}}^T \hat{X} R_{\hat{X}})} \sqrt{\text{Var}(L_{\hat{Y}}^T \hat{Y} R_{\hat{Y}})}} \quad (10)$$

The image datasets  $\hat{X}$  and  $\hat{Y}$  are now transformed to  $P_X = L_{\hat{X}}^T \hat{X} R_{\hat{X}}$  and  $P_Y = L_{\hat{Y}}^T \hat{Y} R_{\hat{Y}}$ .

##### 3.1.2. Reconstruction

In order to perform super-resolution, the LR image  $i_{LR}$  is provided. The LR image is projected to the subspace by

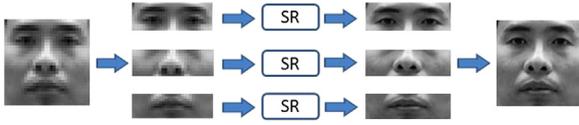
$$P_i^{LR} = L_{\hat{Y}}^T (i_{LR} - \mu_Y) R_{\hat{Y}} \quad (11)$$

We assume that  $P_i^{LR}$  can be reconstructed by a linear combination of its  $K$  nearest neighbors in  $P_Y$  and the coefficients  $w_j$ 's are obtained by minimizing the reconstruction error given by

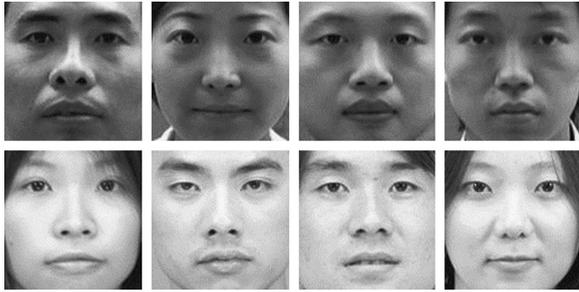
$$\arg\min_{\{w_j\}_{j=1}^K} \left\| P_i^{LR} - \sum_{j=1}^K w_j P_{Y_j} \right\|_F \quad (12)$$

subject to the constraint  $\sum_{j=1}^K w_j = 1$ .  $P_{Y_j}$  denotes a sample in the LR dataset, and  $\|\cdot\|_F$  calculates the Frobenius norm. The details on solving this constrained least square problem can be found in [43].

After obtaining the reconstruction weights  $\{w_j\}_{j=1}^K$ , the projection of the desired HR image  $i_{HR}$  in the 2D CCA space



**Fig. 2.** A face is divided into three parts corresponding to eyes region, nose region, and mouth region. Super-resolution is performed separately for each part and the outputs are merged together to form the high-resolution output.



**Fig. 3.** Sample images from two datasets: CAS-PEAL-R1 dataset [44] (top) and CUHK student dataset [48] (bottom).

is reconstructed by

$$P_i^{HR} = \sum_{j=1}^K w_j P_{X_j} \quad (13)$$

where  $P_{X_j}$  is the HR version corresponding to  $P_{Y_j}$ .

Similar to (11),  $P_i^{HR}$  is related to  $i_{HR}$  by

$$P_i^{HR} = L_X^T (i_{HR} - \mu_X) R_X^\dagger \quad (14)$$

so  $i_{HR}$  is derived as

$$i_{HR} = L_X^{T\dagger} P_i^{HR} R_X^\dagger + \mu_X \quad (15)$$

where  $\dagger$  denotes the Moore–Penrose pseudoinverse operation since  $L_X^T$  and  $R_X$  are not directly invertible.

The super-resolution approach for the whole face mentioned above is based on the rationale that the same neighborhoods are preserved in both HR dataset and LR dataset. Instead of generating a model for the whole face, we divide a face into three parts from top to bottom: eyes part, nose part, and mouth part (for the aligned face images, each part is taken within a predefined region). The same super-resolution procedure is applied directly with the only difference being that the training LR and HR image pairs and the input LR image are now certain parts of the face. The partitioning improves the global reconstruction precision by refining local reconstruction separately. The final result is obtained by stitching the three independently reconstructed parts together as shown in Fig. 2. We average the pixels on the boundaries from different parts to generate a smooth output.

### 3.2. Detail compensation

During face reconstruction, the projection of the face data into a subspace inevitably loses some information and this is often observed as the lack of high-frequency details.

Furthermore, the neighborhood reconstruction itself is essentially an averaging process which further smooths the reconstruction results. To alleviate this problem, we add a detail compensation step in order to generate faces with high-frequency details.

For a HR image  $x_j$  in the training set  $X$ , a Gaussian filter is applied, effectively as a low-pass filter. The output  $\tilde{x}_j$  is a blurred version of the original image that mainly contains the low-frequency (LF) components. By subtracting the low-passed image  $\tilde{x}_j$  from the original image  $x_j$ , an image  $h_j$  that contains mainly high-frequency (HF) components is generated.

The reconstruction weights are computed during the training as in (12) to generate a HF compensation mask  $i_{comp}$  by

$$i_{comp} = \sum_{j=1}^K w_j h_j \quad (16)$$

where  $\{h_j\}_{j=1}^K$  corresponds to the same neighborhood as that of  $\{P_{X_j}\}_{j=1}^K$  and  $\{P_{Y_j}\}_{j=1}^K$ , and  $K$  is the number of chosen nearest neighbors. The final output is given by

$$i_{out} = i_{HR} + i_{comp} \quad (17)$$

Similar as it was done in face reconstruction, each HF image is divided into three parts and the three sets of calculated weights are applied to generate the three reconstructed HF masks. The HF masks are then combined together to form a detail compensation mask for the whole face.

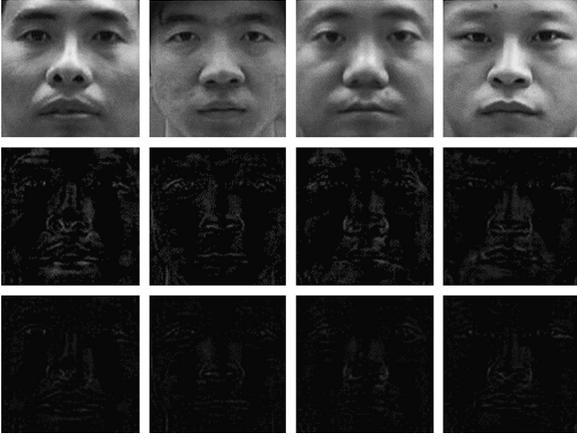
## 4. Experiments

### 4.1. Experimental protocols

#### 4.1.1. Face datasets

We evaluate our method on the CAS-PEAL-R1 dataset [44] which contains face images of 1040 individuals. We use frontal face images of these individuals with normal lighting and neutral expression. All images are cropped and geometrically normalized with the locations of the eyes and mouths fixed to form the HR images of size  $120 \times 120$  and the LR images of size  $30 \times 30$ . Thus, the magnification factor is 4 in our experiments. We selected images of 940 individuals for training and the rest of the images are used for testing. As studied and pointed out in [14] as well as found out empirically in our experiments, the larger the neighborhood size  $K$  is, the better the results of neighborhood reconstruction are. Therefore, in our experiments all the training images are used in the  $K$  nearest neighbor reconstruction. To generate the LF images, a Gaussian filter with  $\sigma = 0.8$  is applied with a  $5 \times 5$  mask. This low-pass filtering of the image simulates the formation of the LR image from the HR image. The selection of the parameters for the Gaussian filter is similar to the settings in the previous work for image SR [45–47].

In addition, we also use CUHK student dataset [48] which contains 188 subjects with frontal faces to test our algorithm. The images in this dataset are cropped and aligned in the same manner. Fig. 3 shows some sample images from both datasets.



**Fig. 4.** Original HR images (top), residue between the original images and whole face based SR results (middle), residue between the original images and part based SR results (bottom). *Best viewed on screen.*

For training, the HR images are divided into three parts of size  $40 \times 120$ ,  $48 \times 120$ , and  $40 \times 120$ . The middle part has 4 rows of pixels at top and at bottom overlapping with the upper part and the lower part. The corresponding LR images are divided to  $10 \times 30$ ,  $12 \times 30$ , and  $10 \times 30$ . For testing, we concatenate the super-resolved parts and average the overlapping pixels. By averaging the overlapping regions the final output is smooth and consistent (see Fig. 7). We choose  $d_1 = d_2 = 30$  in projection matrices empirically to maintain the reconstruction accuracy in an efficient manner. The projection matrices converge in about 10 iterations during optimization.

#### 4.1.2. Methods compared

We compare our results to four *state-of-the-art* methods: artifact free super-resolution method using iterative curve based interpolation (ICBI) [8], sparse representation based super-resolution (SPR) [10], the position-patch based method (PP) [23], and the 1D CCA based face reconstruction method (1D CCA) [14]. Among these methods, 1D CCA and PP are specially designed to super-resolve face images while ICBI and SPR are the super-resolution algorithms for generic images. For ICBI, SPR and PP, we used the default settings provided by the authors of these papers [8,10,23]. We use the same training set in 1D CCA as in our method [14]. For face reconstruction in PP, the input HR–LR pairs come from our training set. Note that we do not compare with [25] since it is not directly related to image super-resolution and we do not compare with [18] since it is primarily designed for face image-sketch synthesis.

#### 4.1.3. Metrics for quantitative evaluation

We calculate the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [49] scores for the super-resolved face images. In addition, we calculate the distortion measure (DM) [41] that evaluates the distortion from the original image in the frequency domain, in which other image quality measures usually do not work. The DM is

calculated by

$$DM = \int_0^{f_{max}} \left[ 1 - DTF \left( \frac{f_r}{f_N} \right) \right] CSF(f_r) df_r \quad (18)$$

where  $f_r$  is the radial frequency,  $f_N$  is the Nyquist frequency and  $f_{max}$  is a predefined value. DTF is a distortion transfer function and CSF is a contrast sensitivity function which approximates the human visual system (HVS).

Furthermore, we also apply a recently introduced SVD-based quality measure [40] in our experiments. The SVD-based image quality measure tries to mimic a human viewer by measuring different distortion types at different levels. In this metric first a graphical measure is calculated for each image block of size  $n \times n$  by

$$D_i = \sqrt{\sum_{i=1}^n (s_i - \hat{s}_i)^2} \quad (19)$$

where  $s_i$  is the singular value of the original block and  $\hat{s}_i$  is the singular value of the distorted block. The global measure is obtained by

$$M_{SVD} = \frac{\sum_{i=1}^{k/n \times k/n} |D_i - D_{mid}|}{\frac{k}{n} \times \frac{k}{n}} \quad (20)$$

where  $D_{mid}$  is the median of the sorted  $D_i$  and  $k \times k$  is the size of the image.

## 4.2. Experimental results

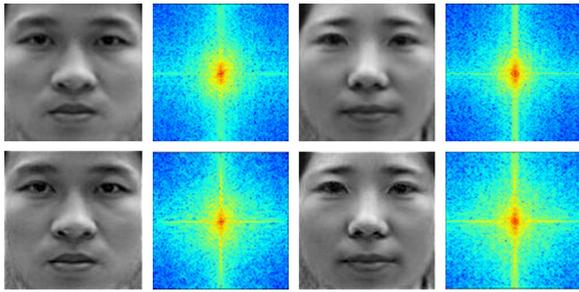
### 4.2.1. Effect of part-based SR

As aforementioned, instead of performing SR on the whole face, we divide the face into three parts and reconstruct each part individually. The final output merges the three parts together. Fig. 4 shows the difference by subtracting the reconstructed faces from the original HR images for some sample images using the holistic method and the part-based method.

Examining the difference between the original images and the proposed part based SR results, we find that reconstructing the entire face using one trained model brings more error especially in the regions of eyes, nose and mouth. By specializing the trained model for a specific part of a face, the output is closer to the ground-truth with less distortions.

### 4.2.2. Effect of detail compensation

In Fig. 5 we examine the Fourier transform on the face images after face reconstruction and detail compensation. The magnitude of the Fourier transform is drawn as the heat map where the magnitude decreases from red to blue. After face reconstruction, the low-frequency components are dominating, as can be seen in the center of the heat map. The magnitude decreases from the center toward the corners of the heat map. This agrees with our visual impression that the reconstructed faces are not sufficiently sharp. After detail compensation, the magnitude is increased in the high-frequency components. Thus, the face images look sharper with less blurriness and more details.



**Fig. 5.** Effects of detail compensation. (Top) Reconstructed faces by 2D CCA. (Bottom) Results after detail compensation. The heat map to the right of the image shows the magnitude of its Fourier transform. *Best viewed on screen.* (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

**Table 1**

Evaluation of the effects of part-based SR and detail compensation. The evaluation metrics include PSNR, SSIM [49], DM [41] and SVD [40]. For all the metrics, the higher score is better.

Metric	Holistic face	Part-based	Detail comp.
PSNR (dB)	32.45	34.46	34.89
SSIM	0.843	0.884	0.885
DM (dB)	35.38	35.87	38.25
SVD	0.667	0.672	0.668

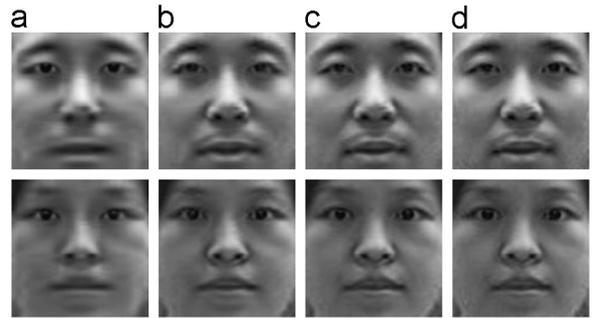
Table 1 shows the effect of part-based SR and detail compensation. The performance gain occurs from holistic face SR to part-based SR for all the three evaluation metrics (PSNR, SSIM, and DM). Further improvement is achieved by detail compensation. The scores of SVD [40] are similar in the comparisons. The quantitative results indicate that both the part-based SR and detail compensation help to improve the output image quality. These quantitative results correspond to the visual observations from Figs. 4 and 5.

#### 4.2.3. Effect of projection dimension

The effect of the dimension  $d_1$  and  $d_2$  in the projection matrices is examined. As shown in Fig. 6, when  $d_1$  and  $d_2$  are small, the reconstruction is not accurate. The choice of  $d_1 = d_2 = 30$  generates good results. As  $d_1$  and  $d_2$  become much larger, the differences in the outputs are not easily noticeable while the computation and memory expenses increase. As a result, we choose  $d_1 = d_2 = 30$  in our experiments as the dimension of the left and right projection matrices.

#### 4.2.4. Comparison with 1D CCA based method

In the 1D CCA based method [14] the images are first converted into vectors. From Fig. 7(e) we can see that although 1D CCA is able to super-resolve the faces, the output images suffer from distortions. This inaccurate reconstruction visually causes the super-resolved image to deviate from the ground-truth image. In other words, the output face images look different from the actual subjects (see the distortions on the subjects' noses, eyes, mouths, and chins). This may degrade the performance of latter processing steps such as face recognition (see



**Fig. 6.** Effects of dimension of the projection matrices. (a)  $d=10$ . (b)  $d=20$ . (c)  $d=30$ . (d)  $d=40$ . *Best viewed on screen.*

Table 2). The reason for the distortions is that the images are first reshaped into 1D vectors and then the relationships in the 1D CCA subspace are explored. However, since the data are intrinsically 2D structured, this reshaping process would inevitably discard the 2D spatial information in the original data representation.

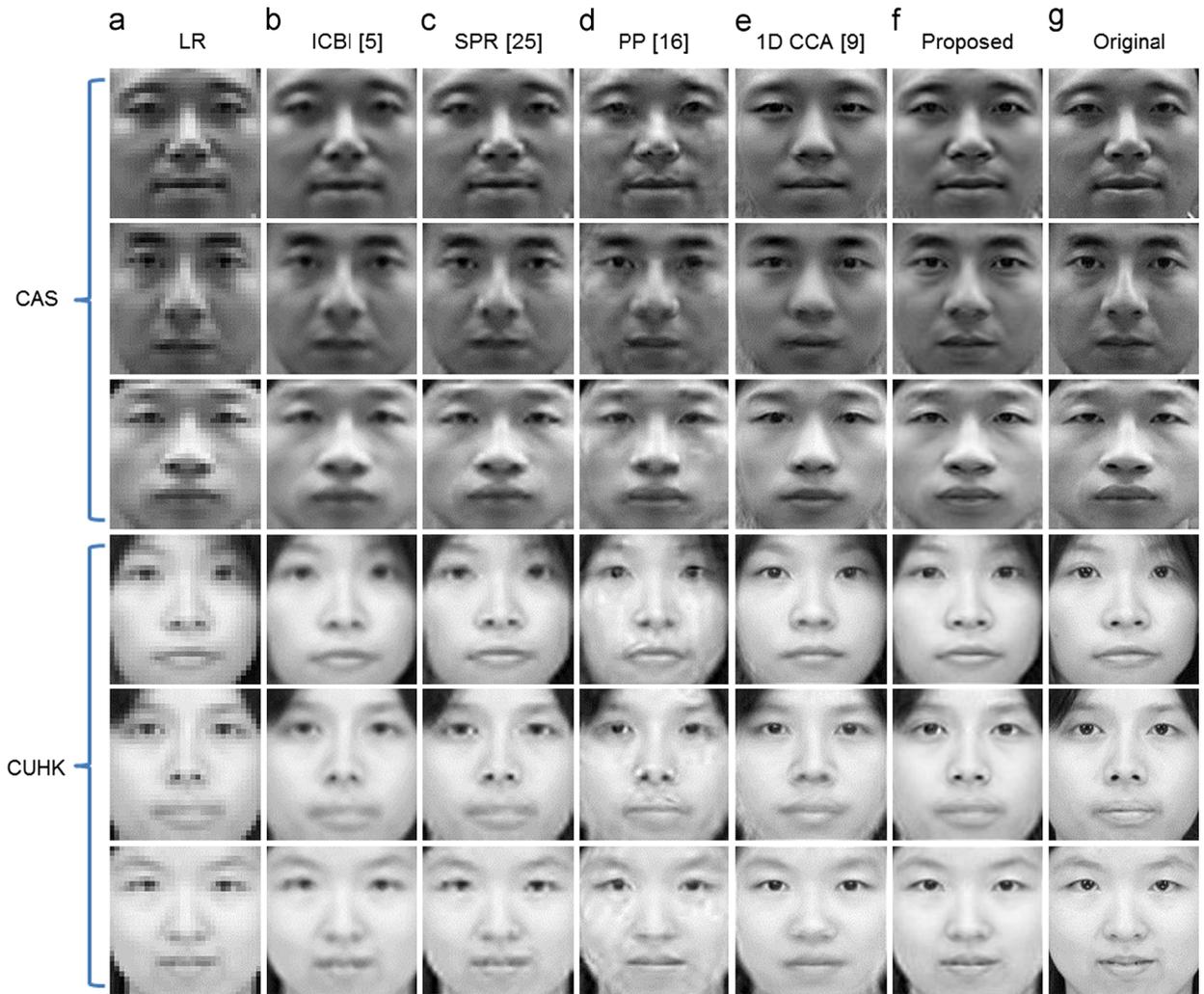
In the proposed 2D CCA based methods, those distortions are significantly reduced (see Fig. 7(f)) with respect to the ground-truth (see Fig. 7(g)). It is evident that by bypassing the image vectorization, the output image is better reconstructed in terms of its underlying structure.

Fig. 8 shows the box plots for the quality measures by PSNR, SSIM, DM, and SVD. Box plot is a non-parametric display of differences between groups of numerical data. The proposed method outperforms the 1D CCA based method in all of the reference based and non-reference based metrics. Especially, 1D CCA yields poor scores for SSIM as a similarity measure, since its results are distorted from the ground-truth. Compared with 1D CCA based method, both the reconstruction error and artifacts are reduced by our method, thus leading to better quantitative scores. In [14] the 1D CCA based method outperforms some of the representative face super-resolution methods including [15], a common baseline method for face super-resolution. For this paper, the proposed method outperforms the method in [14] (see Fig. 8). Therefore, the proposed method is better than [15].

#### 4.2.5. Comparison with other methods

We compare the proposed approach with some *state-of-the-art* methods. Fig. 7 shows sample results with different methods. We summarize the findings as follows:

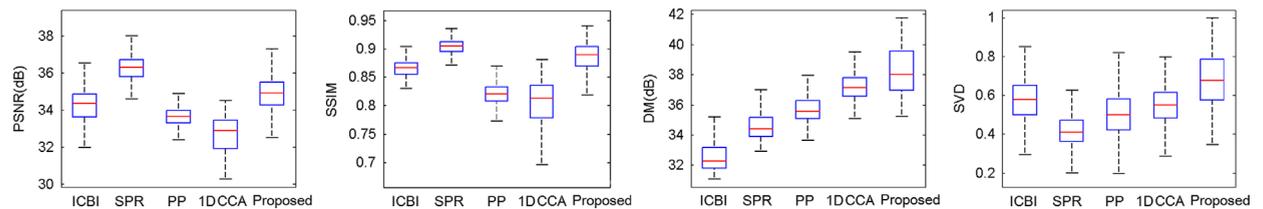
- The results by ICBI [8] (see Fig. 7(b)) do not contain sufficient details and the blurriness in the output is not removed. The interpolation based method is not able to reconstruct facial details.
- SPR based method [10] tackles the SR problem from the perspective of compressed sensing. It is based on the assumption that the sparse representation can be recovered correctly from the downsampled signal. As shown in Fig. 7(c), the results contain more detail and the faces are reconstructed properly. However, the staircase noise is noticeable along the curved edges.



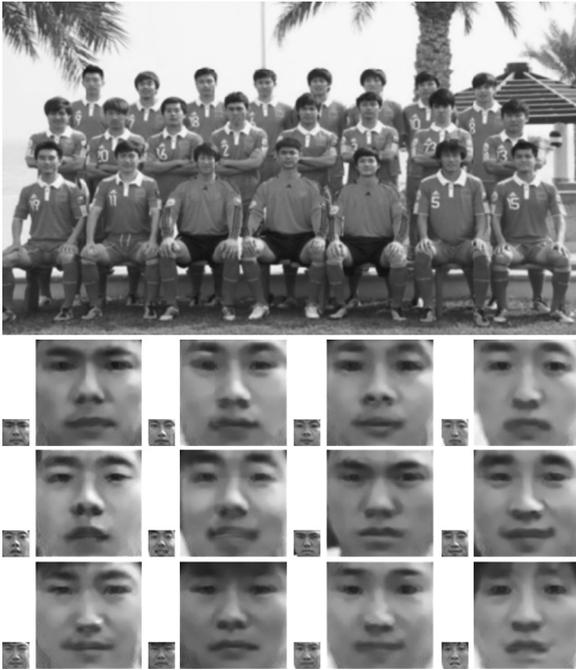
**Fig. 7.** The super-resolution results. Top three rows are from CAS-PEAL-R1 dataset [44] and bottom three rows are from CUHK dataset [48]. (a) Low-resolution images (enlarged by pixel replication). (b) Results by ICBI [8]. (c) Results by SPR [10]. (d) Results by PP [23]. (e) Results by 1D CCA [14]. (f) Results by the proposed method. (g) Original high-resolution images. *Best viewed on screen.*

**Table 2**  
Recognition accuracy using super-resolved images.

Method	ICBI [8]	SPR [10]	PP [23]	1D CCA [14]	Step 1	Proposed
Accuracy (%)	75	97.22	87.5	64.58	98.96	99.31



**Fig. 8.** The box plots of the proposed method and four *state-of-the-art* methods: ICBI [8], SPR [10], PP [23], and 1D CCA [14]. The red bar indicate the median of the results. Above median is the upper quartile and maximum value, while below median is the lower quartile and minimum value. Metrics from left to right: PSNR, SSIM [49], DM [41] and SVD [40]. For all the metrics, the higher score is better. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)



**Fig. 9.** Results on a real world image. (Top) Original image. (Bottom) Some extracted LR faces (small images) and the super-resolved faces (large images). Best viewed on screen.

- PP based method [23] divides the face image into many small connected patches and then uses neighbor embedding to super-resolve each patch separately. The final results are constructed by stitching the small patches together. As can be seen from the results (see Fig. 7(d)), the general structure of the faces is well maintained. The blockiness artifacts are explicitly visible due to the reconstruction in a local manner without the global refinement.
- The 2D spatial structure of the LR faces is well maintained in the 2D CCA subspace since the image data are fed directly in the optimization process without being converted into vectors. Fig. 7(f) shows the final results by our method. In this case, there is more resemblance between the output and the original images. The detail compensation further improves the results by adding more details to the faces, i.e. the contours of the eyes and mouths. Compared to the original HR face images, the final outputs of the proposed algorithm are realistic without explicit artifacts, confirming that the 2D structural information is well maintained by using the 2D CCA method.

The training images are from the CAS dataset only. When applied the trained model to the CUHK dataset, which is taken under different illumination conditions (see Fig. 3), the outputs are still satisfactory. Thus, it is evident that the trained model is not dataset dependent and this merit makes our method generalizable.

The proposed method outperforms PP [23], and ICBI [8] in PSNR and SSIM. For PSNR, the results by SPR [10] are higher than the proposed method. However, this contradicts with our visual examination. In fact, as indicated in [49], PSNR itself does not translate the visual quality to scores faithfully. For SSIM, our method is comparable to

SPR [10]. For DM and SVD, our method yields the best results, which means in these aspects that the proposed method is able to generate highest quality outputs with minimum artifacts and distortion. The results by DM and SVD are more coherent with the visual quality on the super-resolved images from Fig. 7.

#### 4.2.6. Results on real world data

We test the proposed method on a real world image with LR faces. We manually extracted the faces from the picture and aligned them as in the previous experiments. Some samples of the super-resolved face images are shown in Fig. 9. These subjects are not present in the database we used for training and testing above. Note that the quality of the input image is significantly worse than the quality of the images from CAS-PEAL-R1 or CUHK datasets due to noise, blurriness and artifacts caused by compression. Still our algorithm is able to generate reasonably good results.

### 4.3. Effect of super-resolution on recognition

In order to evaluate the effects of super-resolution to face recognition, we conduct a recognition experiment with LBP-based face recognition [50] as the baseline method. The 100 subjects from CAS-PEAL-R1 dataset [44] and 188 subjects from CUHK dataset [48] used for testing in the above experiments are now combined to form a dataset of 288 subjects. The gallery set contains HR image for those subjects and the query data are the super-resolved images using different methods. Table 2 shows the recognition accuracy.

From the recognition rates we can see that the images generated by the proposed methods lead to better recognition result compared to the other methods. The result by 1D CCA [14] is less competitive since the generated faces are distorted from the ground-truth, which adversely affects the recognition performance. The recognition rate using 2D CCA only (step 1) without detail compensation is 98.96%, which is better compared to other methods. By further compensating the details, the highest recognition rate of 99.31% is achieved.

We would expect that by using more sophisticated features and classifiers the recognition rates using images by different super-resolution methods would also increase. However, given parameter settings in the baseline face recognition method, the 2D CCA super-resolved faces result in the best recognition rate. This is in agreement with the visual quality assessment in Fig. 7 in which the faces are faithfully reconstructed using 2D CCA.

### 4.4. Computational complexity

One of the advantages of our method is the simplicity of computation. The computational complexity depends on the solution to solve the generalized eigenvalue problem to solve (8) and (9). Many eigenvalue solvers can be used. For instance, the Arnoldi iteration [51] is an efficient and a popular algorithm. Given a generalized eigenvalue problem  $Ax = \lambda Bx$ , suppose the matrix size of A and B be  $N$  by  $N$ , then the computational complexity using the Arnoldi iteration is  $O(dN^2 + d^2N)$ , where  $d$  is the number of significant eigenvalues. In (8) and (9) for the solution of

**Table 3**

Comparison of training time (in seconds) and average time to super-resolve a face image.

Method	ICBI [8]	SPR [10]	PP [23]	1D CCA [14]	Proposed
Training	–	351.86 s	–	103.36 s	64.30 s
Testing	0.087 s	39.23 s	112.81 s	0.48 s	1.38 s

2D CCA, the sizes of covariance matrices corresponding to  $A$  and  $B$  are small (i.e., they are of the order of image width or height). This implies smaller  $N$  in  $O(dN^2 + d^2N)$ , thus 2D CCA is computationally efficient. On the other hand, in 1D CCA, since images are first reshaped to vectors, the corresponding covariance matrices in (7) are much larger (i.e., they are of the order of total number of pixels in an image), which lead to more expensive 1D CCA as compared to the 2D CCA.

Among the methods compared in this paper, SPR [10], 1D CCA [14], and the proposed 2D CCA based method require a training process. Note that in PP [23] the HR–LR pairs are used to reconstruct the testing image and no explicit model or subspace is formed via training, thus, we do not consider it here as a learning based approach that requires training. All the programs are implemented in MATLAB and were executed on a desktop with a 2.4 GHz CPU and 3 GB of RAM. The implementation of our method is not optimized. Table 3 shows the training time and the average time to super-resolve a face image on CAS-PEAL-R1 and CUHK datasets.

Compared to the other methods involving a training process, our method took significantly less time for training. It is due to the small size of matrices involved in the 2D CCA computation. To super-resolve the LR face image, although ICBI [8] and 1D CCA [14] spent less time, our method is able to generate better results while also keeping the computation time to a small value. When comparing to SPR [10] and PP [23], our method requires much less time to super-resolve an image.

#### 4.5. Discussion

The high quality of the super-resolved image demands accurate alignment in the preprocessing step. The images for training and testing need to be aligned in the same manner (with the positions of eyes and the center of the mouth fixed). Without proper alignment, the quality of the output would degrade. However, this constraint also holds for CCA based method [14] and position-patch based method [23].

#### 5. Conclusions

In this paper, a two-step approach for face super-resolution based on 2D canonical correlation analysis is proposed. One major merit of the proposed method is that our method works directly on the original 2D representation of the image data without converting the images into vectors as it is commonly done in the previous work. This important methodology maintains the intrinsic 2D structure of the face images. Experimental results show that compared to the *state-of-the-art* methods, the super-

resolved faces by the proposed approach are visually realistic and very close to the ground-truth. Various image quality metrics also support that the results by our method are superior to the other methods. The super-resolved images are tested in the recognition task and the results suggest that the super-resolved images by the proposed method achieve the highest accuracy. Due to the small matrices involved in our method, the computation in both training and testing processes is very efficient.

#### Acknowledgements

This work was supported in part by NSF grant 0905671.

#### References

- [1] P. Hennings-Yeomans, S. Baker, B. Kumar, Simultaneous super-resolution and feature extraction for recognition of low-resolution faces, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8. <http://dx.doi.org/10.1109/CVPR.2008.4587810>.
- [2] S.C. Park, M.K. Park, M.G. Kang, Super-resolution image reconstruction: a technical overview, IEEE Signal Processing Magazine 20 (3) (2003) 21–36.
- [3] M. Elad, A. Feuer, Super-resolution reconstruction of image sequences, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (9) (1999) 817–834.
- [4] R. Schultz, R. Stevenson, Extraction of high-resolution frames from video sequences, IEEE Transactions on Image Processing 5 (6) (1996) 996–1011.
- [5] A. Zomet, A. Rav-Acha, S. Peleg, Robust super-resolution, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, 2001, pp. I-645–I-650.
- [6] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, A. Katsaggelos, Softcuts: A soft edge smoothness prior for color image super-resolution, IEEE Transactions on Image Processing 18 (5) (2009) 969–981.
- [7] H. Takeda, S. Farsiu, P. Milanfar, Kernel regression for image processing and reconstruction, IEEE Transactions on Image Processing 16 (2) (2007) 349–366.
- [8] A. Giachetti, N. Asuni, Real time artifact-free image upscaling, IEEE Transactions on Image Processing PP (99) (2011) 1.
- [9] W.T. Freeman, E.C. Pasztor, O.T. Carmichael, Learning low-level vision, International Journal of Computer Vision 40 (2000) 25–47.
- [10] J. Yang, J. Wright, T. Huang, Y. Ma, Image super-resolution as sparse representation of raw image patches, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.
- [11] M. Tappen, C. Liu, A Bayesian approach to alignment-based image hallucination, in: European Conference on Computer Vision, 2012, pp. 236–249.
- [12] X. Gao, K. Zhang, D. Tao, X. Li, Joint learning for single-image super-resolution via a coupled constraint, IEEE Transactions on Image Processing 21 (2) (2012) 469–480.
- [13] X. Gao, K. Zhang, D. Tao, X. Li, Image super-resolution with sparse neighbor embedding, IEEE Transactions on Image Processing 21 (7) (2012) 3194–3205.
- [14] H. Huang, H. He, X. Fan, J. Zhang, Super-resolution of human face image using canonical correlation analysis, Pattern Recognition 43 (7) (2010) 2532–2543.
- [15] C. Liu, H.-Y. Shum, C.-S. Zhang, A two-step approach to hallucinating faces: global parametric model and local nonparametric model, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, 2001, pp. I-192–I-198.
- [16] W. Liu, D. Lin, X. Tang, Hallucinating faces: tensorpatch super-resolution and coupled residue compensation, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, 2005, pp. 478–484.
- [17] Y. Zhuang, J. Zhang, F. Wu, Hallucinating faces: LPH super-resolution and neighbor reconstruction for residue compensation, Pattern Recognition 40 (11) (2007) 3178–3194.
- [18] S. Wang, L. Zhang, Y. Liang, Q. Pan, Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis, in: IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 2216–2223.

- [19] H. Chang, D.-Y. Yeung, Y. Xiong, Super-resolution through neighbor embedding, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, 2004, pp. 1-275–1-282.
- [20] X. He, S. Yan, Y. Hu, P. Niyogi, H.-J. Zhang, Face recognition using Laplacian faces, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (3) (2005) 328–340.
- [21] B. Kumar, R. Aravind, Face hallucination using OLPP and kernel ridge regression, in: IEEE International Conference on Image Processing, 2008, pp. 353–356.
- [22] X. Wang, X. Tang, Hallucinating face by eigentransformation, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 35 (3) (2005) 425–434.
- [23] X. Ma, J. Zhang, C. Qi, Hallucinating face by position-patch, Pattern Recognition 43 (2010) 2224–2236.
- [24] H. Huang, N. Wu, Fast facial image super-resolution via local linear transformations for resource-limited applications, IEEE Transactions on Circuits and Systems for Video Technology 21 (10) (2011) 1363–1377.
- [25] K. Nguyen, S. Sridharan, S. Denman, C. Fookes, Feature-domain super-resolution framework for gabor-based face and iris recognition, in: IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 2642–2649.
- [26] H. Hotelling, Relations between two sets of variates, Biometrika 28 (3/4) (1936) 321–377.
- [27] D.R. Hardoon, S.R. Szedmak, J.R. Shawe-taylor, Canonical correlation analysis: an overview with application to learning methods, Neural Computation 16 (2004) 2639–2664.
- [28] T. Sun, S. Chen, Locality preserving CCA with applications to data visualization and pose estimation, Image and Vision Computing 25 (5) (2007) 531–543.
- [29] Z. Gou, C. Fyfe, A canonical correlation neural network for multicollinearity and functional data, Neural Networks 17 (2) (2004) 285–293.
- [30] W. Yang, D. Yi, Z. Lei, J. Sang, S. Li, 2D–3D face matching using CCA, in: IEEE International Conference on Automatic Face Gesture Recognition, 2008, pp. 1–6.
- [31] W. Zheng, X. Zhou, C. Zou, L. Zhao, Facial expression recognition using kernel canonical correlation analysis (KCCA), IEEE Transactions on Neural Networks 17 (1) (2006) 233–238.
- [32] X. Xu, Z. Mu, Feature fusion method based on KCCA for ear and profile face based multimodal recognition, in: IEEE International Conference on Automation and Logistics, 2007, pp. 620–623.
- [33] N. Correa, T. Adali, Y.-O. Li, V. Calhoun, Canonical correlation analysis for data fusion and group inferences, IEEE Signal Processing Magazine 27 (4) (2010) 39–50.
- [34] A. Nielsen, Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data, IEEE Transactions on Image Processing 11 (3) (2002) 293–305.
- [35] Q.-S. Sun, S.-G. Zeng, P.-A. Heng, D.-S. Xia, Feature fusion method based on canonical correlation analysis and handwritten character recognition, in: Control, Automation, Robotics and Vision Conference, vol. 2, 2004, pp. 1547–1552.
- [36] A. Li, S. Shan, X. Chen, W. Gao, Maximizing intra-individual correlations for face recognition across pose differences, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 605–611.
- [37] L. An, N. Thakoor, B. Bhanu, Vehicle logo super-resolution by canonical correlation analysis, in: IEEE International Conference on Image Processing, 2012.
- [38] S.H. Lee, S. Choi, Two-dimensional canonical correlation analysis, IEEE Signal Processing Letters 14 (10) (2007) 735–738.
- [39] H. Wang, Local two-dimensional canonical correlation analysis, IEEE Signal Processing Letters 17 (11) (2010) 921–924.
- [40] A. Shnayderman, A. Gusev, A. Eskicioglu, An SVD-based grayscale image quality measure for local and global assessment, IEEE Transactions on Image Processing 15 (2) (2006) 422–429.
- [41] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans, A. Bovik, Image quality assessment based on a degradation model, IEEE Transactions on Image Processing 9 (4) (2000) 636–650.
- [42] J. Yang, D. Zhang, A. Frangi, J. Yu Yang, Two-dimensional PCA: a new approach to appearance-based face representation and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (1) (2004) 131–137.
- [43] L.K. Saul, S.T. Roweis, Think globally, fit locally: unsupervised learning of low dimensional manifolds, Journal of Machine Learning Research 4 (2003) 119–155.
- [44] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, D. Zhao, The CAS-PEAL large-scale Chinese face database and baseline evaluations, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 38 (1) (2008) 149–161.
- [45] W. Dong, D. Zhang, G. Shi, X. Wu, Nonlocal back-projection for adaptive image enlargement, in: IEEE International Conference on Image Processing, 2009, pp. 349–352. <http://dx.doi.org/10.1109/ICIP.2009.5414423>.
- [46] C.-Y. Yang, J.-B. Huang, M.-H. Yang, Exploiting self-similarities for single frame super-resolution, in: Asian Conference on Computer Vision, 2010, pp. 1807–1818.
- [47] L. An, B. Bhanu, Improved image super-resolution by support vector regression, in: IEEE International Joint Conference on Neural Networks, 2011, pp. 696–700.
- [48] X. Wang, X. Tang, Face photo-sketch synthesis and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 31 (11) (2009) 1955–1967.
- [49] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing 13 (4) (2004) 600–612.
- [50] T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: application to face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (12) (2006) 2037–2041.
- [51] L.N. Trefethen, D. Bau, Numerical Linear Algebra, SIAM: Society for Industrial and Applied Mathematics, 1997.